Revisiting 2x2 matrix optics: Complex vectors, Fermion combinatorics, and Lagrange invariants

Quirino M. Sugon Jr.* and Daniel J. McNamara

Ateneo de Manila University, Department of Physics, Loyola Heights, Quezon City, Philippines 1108

*Also at Manila Observatory, Upper Atmosphere Division, Ateneo de Manila University Campus

e-mail: qsugon@observatory.ph

December 3, 2008

Abstract. We propose that the height-angle ray vector in matrix optics should be complex, based on a geometric algebra analysis. We also propose that the ray's 2×2 matrix operators should be right-acting, so that the matrix product succession would go with light's left-to-right propagation. We express the propagation and refraction operators as a sum of a unit matrix and an imaginary matrix proportional to the Fermion creation or annihilation matrix. In this way, we reduce the products of matrix operators into sums of creation-annihilation product combinations. We classify ABCD optical systems into four: telescopic, inverse Fourier transforming, Fourier transforming, and imaging. We show that each of these systems have a corresponding Lagrange theorem expressed in partial derivatives, and that only the telescopic and imaging systems have Lagrange invariants.

1 Introduction

a. Complex Vectors. In 2×2 matrix optics, a ray is normally described by a column vector as given by Nussbaum and Phillips[1]:

$$\mathbf{r} = \begin{pmatrix} x \\ n\alpha \end{pmatrix} = \mathbf{e}_1 x + \mathbf{e}_2 n\alpha, \tag{1}$$

except that we interchanged the coefficients. This equation is problematic: α is an angle and not a distance. The Cartesian coordinate system is a system for locating a point in space in terms of distances from a fixed point measured along orthogonal lines. But in what space do angles live? In an imaginary vector space?

Yes. To see why this is so, let us first recall the orthonormality axioms in geometric algebra[2]:

$$\mathbf{e}_j^2 = \mathbf{e}_k^2 = 1, \tag{2}$$

$$\mathbf{e}_j\mathbf{e}_k = -\mathbf{e}_k\mathbf{e}_j, \quad j \neq k, \tag{3}$$

where $j, k \in \{1, 2, 3\}$. That is, the square of each unit vector is unity and that the product of two perpendicular vectors anticommute. Notice that the multiplicative inverse of a unit vector is itself.

We know from Yariv[3] that the paraxial angle α is the slope of the function $x = x(z)$:

$$\frac{dx}{dz} = \tan\alpha \approx \alpha. \tag{4}$$

But if we define $\mathbf{z} = z\mathbf{e}_3$ and $\mathbf{x} = x\mathbf{e}_1$, then by the rules of geometric algebra, we have

$$\frac{d\mathbf{x}}{d\mathbf{z}} = \mathbf{e}_3^{-1}\mathbf{e}_1\frac{dx}{dz} = \mathbf{e}_3\mathbf{e}_1\frac{dx}{dz} = i\mathbf{e}_2\alpha, \tag{5}$$

where $i = \mathbf{e}_1\mathbf{e}_2\mathbf{e}_3$ is the unit trivector that behaves like the unit imaginary number[4]. Therefore, instead of Eq. (1), we write

$$\hat{r} = \begin{pmatrix} x \\ n\alpha i \end{pmatrix} = \mathbf{e}_1 x + i\mathbf{e}_2 n\alpha, \tag{6}$$

which is a complex vector like the electromagnetic field $\hat{F} = \mathbf{E} + i\mathbf{B}$[5]. Note that we adopted the convention that lengths like height and radius are dimensionless. (Alternatively, we may replace x by ζx, where ζ is a unit quantity with dimension of inverse length).

b. Fermion Combinatorics. In most matrix optics texts, the convention is light travelling left to right. Yet the left-acting propagation and refraction matrix operators used are multiplied from right to left:

$$\mathbf{r}' = \mathsf{M}'_N\mathsf{M}'_{N-1}\cdots\mathsf{M}'_2\mathsf{M}'_1\mathbf{r}. \tag{7}$$

So we propose a more logical way: define the matrix operators to be right-acting[6], so that they multiply from left to right, in the same direction of the light's propagation. That is,

$$\mathbf{r}' = \mathbf{r}\mathsf{M}_1\mathsf{M}_2\cdots\mathsf{M}_{N-1}\mathsf{M}_N. \tag{8}$$

Here, the action of the right-acting matrix is defined by the action of its left-acting transpose, as given in Symon[7]:

$$\mathsf{M}^T \cdot \mathbf{r} = \mathbf{r} \cdot \mathsf{M}. \tag{9}$$

Because the ray operators are 2×2 matrices, we may decompose them as a linear combination of single-element, unit matrices, as done by Campbell[8] and Harris[9]:

$$\mathsf{M} = M_{11}\mathsf{e}_{11} + M_{12}\mathsf{e}_{12} + M_{21}\mathsf{e}_{21} + M_{22}\mathsf{e}_{22}, \tag{10}$$

where

$$\mathsf{e}_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{e}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{11}$$

$$\mathsf{e}_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \mathsf{e}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{12}$$

are the Fermion matrices in Sakurai[10] and Le Bellac[11]. The dyadics e_{12} and e_{21} may represent either the creation operator $\hat{a}^\dagger$ or the annihilation operator $\hat{a}$, depending on the column matrix representations of $\mathbf{e}_1$ and $\mathbf{e}_2$.

For the ray vector $\mathbf{r} = x\mathbf{e}_1 + n\alpha\mathbf{e}_2$, the left-acting propagation and refraction matrices are given in Klein and Furtak (transposed matrices)[12]:

$$\mathsf{T} = 1 + \mathsf{e}_{12}D/n, \tag{13}$$

$$\mathsf{R} = 1 - \mathsf{e}_{21}P, \tag{14}$$

where 1 is the unit matrix. Thus, a lens system becomes a product of propagation and refraction matrices:

$$\begin{aligned} \mathsf{M} &= \mathsf{T}_n\mathsf{R}_n \cdots \mathsf{T}_2\mathsf{R}_2\mathsf{T}_1\mathsf{R}_1 \\ &= \prod_{k=1}^{n}(1 + \mathsf{e}_{12}D_k/n_k)(1 + \mathsf{e}_{21}P_k), \end{aligned} \tag{15}$$

Later, we shall recast these equations for our complex ray vectors and right-acting matrices. We shall also present new methods for computing the system matrix M. These methods are based on the Fermion identies satisfied by the four dyadic operators e_{11}, e_{12}, e_{21}, and e_{22}. In particular, we shall study the allowed combinations of $\mathsf{e}_{12}D_k/n_k$ and $\mathsf{e}_{12}P_{k'}$ and determine the matrix component basis $\mathsf{e}_{kk'}$ of the product chain.

c. Lagrange Invariants. The Lagrange theorem or the Smith-Helmholtz relationship[13, 14] is stated by Welford[15] as

$$nu\eta = n'u'\eta', \tag{16}$$

where n is refractive index of the input medium, the object height, u is the angle subtended from the object, and η is the height of the object; their primed counterparts correspond to those of the image. The quantity $nu\eta$ is called the Lagrange invariant. Later, using our x and $n\alpha$ variables, we shall show that we may recast Lagrange's theorem in Eq. (16) in terms of partial derivatives:

$$\frac{x'}{x}\frac{\partial(n'\alpha')}{\partial(n\alpha)} = -1; \quad x, x' = \text{constants}. \tag{17}$$

The quantity $|x\delta(n\alpha)|$ is the Lagrange invariant.

d. Outline. We shall divide the paper into five sections. The first section is Introduction. In the second section, we discuss the algebra of right-acting matrices and their actions on column vectors. In the third section, we shall introduce the complex ray vector and its right-acting propagation and refraction matrices. We shall use the properties of Fermion creation-annihilation matrices to compute the system matrices of thin and thick lenses. In the fourth section, we shall revisit the classification of optical systems: telescopic, Fourier transforming, inverse Fourier transforming, and imaging. We shall derive their Lagrange theorems and see if we can define their corresponding Lagrange invariants. We shall also rederive the Moebius transform and the Newton's equation for the imaging system. The fifth section is Conclusions.

2 Matrix Algebra

Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be two orthonormal vectors represented as column matrices,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{18}$$

and let e_{11}, e_{12}, e_{21}, and e_{22} be the four Fermion matrices in Eqs. (11) and (12). The left and right action of the dyadic operators on $\mathbf{e}_1$ and $\mathbf{e}_2$ are defined by the following relations[16]:

$$\mathbf{e}_\lambda \cdot \mathsf{e}_{\mu\nu} = \delta_{\lambda\mu}\mathbf{e}_\nu, \tag{19}$$

$$\mathsf{e}_{\mu\nu} \cdot \mathbf{e}_\lambda = \delta_{\nu\lambda}\mathbf{e}_\mu, \tag{20}$$

and

$$\cdot\, \mathsf{e}_{\mu'\nu'} \cdot \mathsf{e}_{\mu\nu} = \delta_{\nu'\mu}(\cdot\, \mathsf{e}_{\mu'\nu}), \tag{21}$$

$$\mathsf{e}_{\mu'\nu'} \cdot \mathsf{e}_{\mu\nu}\cdot = \delta_{\nu'\mu}\mathsf{e}_{\mu'\nu}\cdot, \tag{22}$$

where $\lambda, \mu, \nu \in \{1, 2\}$. Notice that we can rederive these relations if we adopt the definitions

$$\cdot\, \mathsf{e}_{\mu\nu} = \cdot\, \mathbf{e}_\mu\mathbf{e}_\nu, \tag{23}$$

$$\mathsf{e}_{\mu\nu}\cdot = \mathbf{e}_\mu\mathbf{e}_\nu\cdot, \tag{24}$$

with the understanding that the dot product takes precedence over the juxtaposition (geometric) product.

Let $\cdot\mathsf{M}$ be a right-acting 2×2 matrix and let $\mathsf{M}^T\cdot$ be its left-acting transpose:

$$\cdot\, \mathsf{M} = M_{11}\mathsf{e}_{11} + M_{12}\mathsf{e}_{12} + M_{21}\mathsf{e}_{21} + M_{22}\mathsf{e}_{22}, \tag{25}$$

$$\mathsf{M}^T\cdot = M_{11}\mathsf{e}_{11} + M_{12}\mathsf{e}_{21} + M_{21}\mathsf{e}_{12} + M_{22}\mathsf{e}_{22}. \tag{26}$$

The action of these two matrices on the vector

$$\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \tag{27}$$

are related by

$$\mathbf{r}' = \mathsf{M}^T \cdot \mathbf{r} = \mathbf{r} \cdot \mathsf{M}, \tag{28}$$

where $\mathbf{r}'$ is another vector. That is,[6]

$$\begin{aligned} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} &= \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \end{aligned} \tag{29}$$

Hence,

$$x_1' = x_1 M_{11} + x_2 M_{21}, \tag{30}$$
$$x_2' = x_1 M_{12} + x_2 M_{22}. \tag{31}$$

Notice that the column-column multiplication in the action of right-acting matrices is simpler than the row-column multiplication in that of left-acting matrices.

3 Matrix Optics

3.1 Ray Cliffor

Let us define the complex height-angle vector $\hat{r}$ as

$$\hat{r} = x\mathbf{e}_1 + n\alpha i\mathbf{e}_2 = \begin{pmatrix} x \\ n\alpha i \end{pmatrix}. \tag{32}$$

Here, the vector $\mathbf{e}_3$ as the optical axis pointing to the right, $\mathbf{e}_1$ as pointing upwards, and $\mathbf{e}_2$ as pointing out of the paper. We define the light ray to be moving from left to right. The height x of the ray is positive if the ray is above the optical axis and negative if below. The angle α is positive if the ray is inclined and negative if declined. (Fig. 1)

To define the paraxial angle α more precisely, we use sign functions[17, 18]. If $\boldsymbol{\sigma}$ is the direction of propagation of the light ray as it moves close to the direction of the optical axis $\mathbf{e}_3$, then the ray's angle of inclination α with respect to $\mathbf{e}_3$ is[19]

$$\alpha = \theta_\sigma \approx c_{\sigma x}\theta_{\sigma z}, \tag{33}$$

where

$$c_{\sigma x} = \frac{\boldsymbol{\sigma} \cdot \mathbf{e}_1}{|\boldsymbol{\sigma} \cdot \mathbf{e}_1|}, \tag{34}$$
$$\theta_{\sigma z} = \cos^{-1}(\boldsymbol{\sigma} \cdot \mathbf{e}_3). \tag{35}$$

The sign function $c_{\sigma x}$ is the relative direction of the $\boldsymbol{\sigma}$ along the axis $\mathbf{e}_1$, with $+1$ meaning along and -1 opposite. The angle $\theta_{\sigma z}$ is magnitude of the angle between $\boldsymbol{\sigma}$ and $\mathbf{e}_3$.

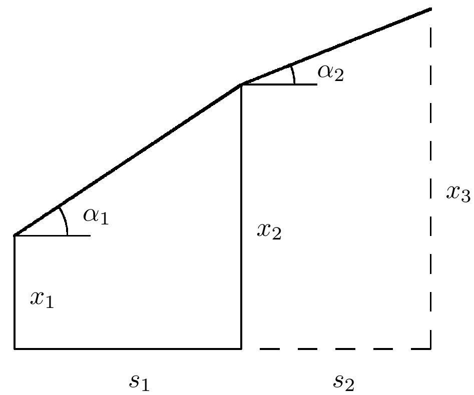

Figure 1: A paraxial ray with height x_1 and inclination angle α_1 moves to the right by a distance s_1 until it hits an refracting surface. The height of the ray becomes x_2 and its inclination angle changes to α_2.

3.2 Matrix Operators

When light propagates, the paraxial meridional angle α remains constant. So the ray tracing equations are

$$x' = x + s\alpha, \tag{36}$$
$$\alpha' = \alpha, \tag{37}$$
$$z' = z + s. \tag{38}$$

In most texts, the last equation is assumed.

Using the definition of the ray cliffor $\hat{r}$ in Eq. (32), Eqs. (36) and (37) may be combined as

$$\hat{r}' = \hat{r} \cdot \mathsf{M}_S = \hat{r} \cdot (1 + \mathsf{S}). \tag{39}$$

where

$$\mathsf{S} = -iS\mathsf{e}_{21} = -i\frac{s}{n}\mathsf{e}_{12}. \tag{40}$$

That is,

$$\begin{pmatrix} x' \\ n'\alpha' i \end{pmatrix} = \begin{pmatrix} x \\ n\alpha i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -is/n & 1 \end{pmatrix}. \tag{41}$$

Except for the $-i$, the propagation matrix is the same as that given by Hecht[20].

On the other hand, when light refracts the height x of the light ray remains constant. So the ray tracing equations are

$$x' = x, \tag{42}$$
$$n'\alpha' = n\alpha - Px, \tag{43}$$
$$z' = z. \tag{44}$$

Here, P is the power of the interface[21]:

$$P = c_{\eta z}\frac{n - n'}{R}, \tag{45}$$

where

$$c_{\eta z} = \frac{\boldsymbol{\eta} \cdot \mathbf{e}_3}{|\boldsymbol{\eta} \cdot \mathbf{e}_3|} \tag{46}$$

is the sign function describing the relative direction of the outward normal vector $\boldsymbol{\eta}$ to the spherical interface of radius R, with respect to the optical axis $\mathbf{e}_3$. If $c_{\eta z} = +1$, $\boldsymbol{\eta}$ is approximately along $\mathbf{e}_3$; if $c_{\eta z} = -1$, $\boldsymbol{\eta}$ is opposite.

Using the definition of the ray cliffor $\hat{r}$ in Eq. (32), Eqs. (42) and (43) may be combined into

$$\hat{r}' = \hat{r} \cdot \mathsf{M}_P = \hat{r} \cdot (1 + \mathsf{P}), \tag{47}$$

where

$$\mathsf{P} = -iP\mathsf{e}_{12}. \tag{48}$$

That is,

$$\begin{pmatrix} x' \\ n'\alpha' i \end{pmatrix} = \begin{pmatrix} x \\ n\alpha i \end{pmatrix} \begin{pmatrix} 1 & -iP \\ 0 & 1 \end{pmatrix}. \tag{49}$$

Again, except for the $-i$, the refraction matrix is similar to that given by Hecht[20].

We may also rewrite the propagation and refraction matrices by using the definition of the exponential of the matrix A as

$$e^{\mathsf{A}} = 1 + \mathsf{A} + \frac{\mathsf{A}^2}{2!} + \frac{\mathsf{A}^3}{3!} + \ldots, \tag{50}$$

so that if $\mathsf{A}^2 = 0$, then

$$e^{\mathsf{A}} = 1 + \mathsf{A}. \tag{51}$$

Thus, the exponential of a null-square matrix is the sum of the matrix itself and the unit matrix.

Because e_{12} and e_{21} are null-square matrices,

$$\mathsf{e}_{12}^2 = \mathsf{e}_{21}^2 = 0, \tag{52}$$

then we may use Eq. (51) to express the matrices M_S and M_P in Eqs. (39) and (47) as

$$\begin{aligned} \mathsf{M}_S &= \; = e^{\mathsf{S}} = 1 + \mathsf{S} = 1 - iS\mathsf{e}_{21}, & (53)\\ \mathsf{M}_P &= \; = e^{\mathsf{P}} = 1 + \mathsf{P} = 1 - iP\mathsf{e}_{12}. & (54) \end{aligned}$$

Except for the $-i$ factor, these exponential forms of the propagation and refraction operators were used before by Simon and Wolf[22]. These forms are useful when we have a series of propagations or of refractions. For these cases, we only have to add the arguments of the exponentials, as what we shall see next.

3.3 System Matrix

An optical system may be considered a black box: we do not know what is inside. All we know is that, in general, the relationship between the input ray $\hat{r}$ and output ray $\hat{r}'$ is given by

$$\hat{r}' = \hat{r} \cdot \mathsf{M}, \tag{55}$$

where M is a 2×2 matrix. That is,

$$\begin{pmatrix} x' \\ n'\alpha' i \end{pmatrix} = \begin{pmatrix} x \\ n\alpha i \end{pmatrix} \begin{pmatrix} A & -iC \\ -iB & D \end{pmatrix}, \tag{56}$$

so that

$$\begin{aligned} x' &= Ax + Bn\alpha, & (57)\\ n'\alpha' &= -Cx + Dn\alpha. & (58) \end{aligned}$$

Notice that these equations are similar to those in the literature, save for the sign of C.

In general, the system matrix M is a product of propagation and refraction operators (c.f. [23]):

$$\mathsf{M} = \prod_{k=1}^{N} \mathsf{M}_{S_k}\mathsf{M}_{P_k} = \prod_{k=1}^{N} e^{\mathsf{S}_k} e^{\mathsf{P}_k} = \prod_{k=1}^{N} (1 + \mathsf{S}_k)(1 + \mathsf{P}_k), \tag{59}$$

where

$$\begin{aligned} \mathsf{S}_k &= -iS_k\mathsf{e}_{21} = -i\frac{s_k}{n_k}\mathsf{e}_{21}, & (60)\\ \mathsf{P}_k &= -iP_k\mathsf{e}_{12} = -ic_{\eta k}\frac{n_k - n_{k+1}}{R}\,\mathsf{e}_{12}. & (61) \end{aligned}$$

Note that the determinant of M is unity[24],

$$|\mathsf{M}| = \prod_{k=1}^{N} |e^{\mathsf{S}_k}||e^{\mathsf{P}_k}| = 1, \tag{62}$$

because its factors have a determinant of unity,

$$|e^{\mathsf{S}_k}| = |e^{\mathsf{P}_k}| = 1. \tag{63}$$

Let us take some special cases. If $\mathsf{P}_k = 0$ for all k, then

$$\mathsf{M} = \prod_{k=1}^{N} \mathsf{M}_{Sk} = \prod_{k=1}^{N} e^{\mathsf{S}_k} = e^{\mathsf{S}} = 1 + \mathsf{S}, \tag{64}$$

where

$$\mathsf{S} = \sum_{k=1}^{N} \mathsf{S}_k = -i\mathsf{e}_{21} \sum_{k=1}^{N} S_k. \tag{65}$$

In other words, the reduced distance[13] or the index-normalized path length $S = s/n$ of a sequence of vertical interfaces (parallel to the $xy-$plane) is equal to the sum of the index-normalized path lengths between successive interfaces. (Path length is normally defined as ns.)

On the other hand, if $\mathsf{S}_k = 0$ for all k, then

$$\mathsf{M} = \prod_{k=1}^{N} \mathsf{M}_{Pk} = \prod_{k=1}^{N} e^{\mathsf{P}_k} = e^{\mathsf{P}} = 1 + \mathsf{P}, \tag{66}$$

where

$$\mathsf{P} = \sum_{k=1}^{N} \mathsf{P}_k = -i\mathsf{e}_{12} \sum_{k=1}^{N} P_k. \tag{67}$$

In other words, the total power P of a sequence of refracting surfaces separated by negligible distances is equal to the sum of the individual powers of each interface.

3.4 Fermion Combinatorics

In general, we cannot simply add the arguments of the exponentials in Eq. (59) because

$$\mathsf{S}_k\mathsf{P}_\ell \neq \mathsf{P}_\ell\mathsf{S}_k. \tag{68}$$

for all subscripts k and ℓ. So instead of the product-of-exponentials form, we shall use the binomial product form and study how to expedite its expansion.

To expand the binomial product form in Eq. (59), we note several things, which we shall label as rules:

Rule 1. The only allowed products are those with alternating S and P factors, because

$$\mathsf{S}_k\mathsf{S}_\ell = \mathsf{P}_k\mathsf{P}_\ell = 0. \tag{69}$$

Rule 2. Since matrix multiplication is not generally commutative, then the order of the factors must be preserved. This means that the values of the subscript k must be increasing from left to right; if the subscripts are the same, S_k should come before P_k.

Rule 3. The matrix S_k is an $-i\mathsf{e}_{21}$ quantity; P_k is a $-i\mathsf{e}_{12}$ quantity. Because the allowed products are those with alternating S and P factors, then the Fermion basis of the product depends only on the first and last factors:

$$\mathsf{S}_k \cdots \mathsf{S}_\ell, \qquad (-i)^L \mathsf{e}_{21}, \tag{70}$$
$$\mathsf{S}_k \cdots \mathsf{P}_\ell, \qquad (-i)^L \mathsf{e}_{22}, \tag{71}$$
$$\mathsf{P}_k \cdots \mathsf{S}_\ell, \qquad (-i)^L \mathsf{e}_{11}, \tag{72}$$
$$\mathsf{P}_k \cdots \mathsf{P}_\ell, \qquad (-i)^L \mathsf{e}_{12}, \tag{73}$$

where L is the number of factors in the product chain.

Rule 4. The unit number 1 is a sum of e_{11} and e_{22},

$$1 = \mathsf{e}_{11} + \mathsf{e}_{22}. \tag{74}$$

These four rules let us compute the system matrix in a systematic way, by simply listing down the allowed combinations of S and P.

3.5 Thin and Thick Lenses

To illustrate our four combinatorial rules, let us compute the expansions of M in Eq. (59) for $k = 1$ and $k = 2$.

Case $k = 1$**.** The matrix M is

$$\begin{aligned} \mathsf{M} &= (1 + \mathsf{S}_1)(1 + \mathsf{P}_1) \\ &= 1 + (\mathsf{S}_1 + \mathsf{P}_1) + \mathsf{S}_1\mathsf{P}_1. \end{aligned} \tag{75}$$

In Fermion basis, this is

$$\begin{aligned} \mathsf{M} &= \mathsf{e}_{11}(1) + \mathsf{e}_{12}(-i)P_1 \\ &\quad + \mathsf{e}_{21}(-i)S_1 + \mathsf{e}_{22}(1 + (-i)^2 S_1 P_1). \end{aligned} \tag{76}$$

That is,

$$\mathsf{M} = \begin{pmatrix} 1 & -iP_1 \\ -iS_1 & 1 - S_1P_1 \end{pmatrix}. \tag{77}$$

Case $k = 2$**.** The matrix M is

$$\begin{aligned} \mathsf{M} &= (1 + \mathsf{S}_1)(1 + \mathsf{P}_1)(1 + \mathsf{S}_2)(1 + \mathsf{P}_2) \\ &= 1 + (\mathsf{S}_1 + \mathsf{P}_1 + \mathsf{S}_2 + \mathsf{P}_2) \\ &\quad + (\mathsf{S}_1\mathsf{P}_1 + \mathsf{S}_1\mathsf{P}_2 + \mathsf{S}_2\mathsf{P}_2 + \mathsf{P}_1\mathsf{S}_2) \\ &\quad + (\mathsf{S}_1\mathsf{P}_1\mathsf{S}_2 + \mathsf{P}_1\mathsf{S}_2\mathsf{P}_2) + \mathsf{S}_1\mathsf{P}_1\mathsf{S}_2\mathsf{P}_2. \end{aligned} \tag{78}$$

In Fermion basis, this is

$$\begin{aligned} \mathsf{M} &= \mathsf{e}_{11}(1 + (-i)^2 P_1 S_2) \\ &\quad + \mathsf{e}_{12}(-iP_1 - iP_2 + (-i)^3 P_1 S_2 P_2) \\ &\quad + \mathsf{e}_{21}(-iS_1 - iS_2 + (-i)^3 S_1 P_1 S_2) \\ &\quad + \mathsf{e}_{22}(1 + (-i)^2 S_1 P_1 + (-i)^2 S_1 P_2 + (-i)^2 S_2 P_2 \\ &\qquad + (-i)^4 S_1 P_1 S_2 P_2). \end{aligned} \tag{79}$$

That is,

$$\mathsf{M} = \begin{pmatrix} (1 - P_1S_2) & -i(P_1 + P_2 - P_1S_2P_2) \\ -i\begin{pmatrix} S_1 + S_2 \\ -S_1P_1S_2 \end{pmatrix} & \begin{pmatrix} 1 - S_1P_1 - S_1P_2 - S_2P_2 \\ + S_1P_1S_2P_2 \end{pmatrix} \end{pmatrix} \tag{80}$$

Check. To check our computations, we set $S_1 = 0$ in Eq. (80) to get

$$\mathsf{M}_{thick} = \begin{pmatrix} (1 - P_1S_2) & -i(P_1 + P_2 - P_1S_2P_2) \\ -iS_2 & (1 - S_2P_2) \end{pmatrix}, \tag{81}$$

Notice that except for the $-i$, Eq. (81) is similar in form to the system matrix for a thick lens as given in Klein and Furtak (they use left-acting matrices and angle-height vectors)[25].

Furthermore, if we set $S_2 = 0$ in Eq. (81), then we obtain the corresponding thin lens matrix[26]:

$$\mathsf{M}_{thin} = \begin{pmatrix} 1 & -i(P_1 + P_2) \\ 0 & 1 \end{pmatrix} \tag{82}$$

Comparison of Eq. (82) with the refraction matrix operator M_P in Eq. (49) yields the thin lens power

$$P = P_1 + P_2, \tag{83}$$

which is what we expect. That is, the total power of a thin lens is equal to the sum of the powers of its refracting surfaces.

4 Optical Systems

So far, we have considered the input and output rays to be very close to the optical black box. We shall now relax this restriction.

Let the matrix M_{box} describe a black box:

$$\mathsf{M}_{\text{box}} = \begin{pmatrix} M_{11} & -iM_{22} \\ -iM_{21} & M_{22} \end{pmatrix}, \tag{84}$$

which satisfies Eq. (62),

$$|\mathsf{M}_{\text{box}}| = M_{11}M_{22} + M_{12}M_{21} = 1. \tag{85}$$

Notice that unlike the determinant of standard system matrices with real coefficients, the determinant of M_{box} in Eq. (85) is not a difference but a sum.

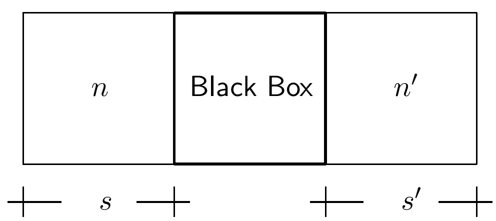

Figure 2: An optical system with a black box. The input side is at the distance s to the left of the box; the output side is at a distance of s' to the right of the box.

To the left of the box at the reduced distance $S = s/n$ is the input ray characterized by height x and optical angle $n\alpha$. To the right of the box at an reduced distance $S' = s'/n'$ is the output ray characterized by height x' and dilated angle $n'\alpha'$. In other words, the system matrix M in Eq. (56) is

$$\begin{aligned} \mathsf{M} &= \begin{pmatrix} A & -iC \\ -iB & D \end{pmatrix} = \mathsf{M}_S \mathsf{M}_{\text{box}} \mathsf{M}_{S'} \\ &= \begin{pmatrix} 1 & 0 \\ -iS & 1 \end{pmatrix} \begin{pmatrix} M_{11} & -iM_{12} \\ -iM_{21} & M_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -iS' & 1 \end{pmatrix} \end{aligned} \tag{86}$$

Except for the $-i$, this matrix product is similar in form to that of Simmons and Guttmann[27]. (In Ghatak[28] and in Nussbaum and Phillips[29] the matrix coefficients of M_{box} are labeled by $-a$, b, c, and $-d$.)

From Eq. (86) we can arrive at two conclusions. First, because the sytem matrix is a product of matrices with unit determinants, then

$$|\mathsf{M}_S| = AB + CD = 1. \tag{87}$$

Second, the elements of the system matrix are given by

$$\begin{aligned} A &= M_{11} - M_{12}S', & (88) \\ B &= M_{21} + M_{22}S' + M_{11}S - M_{12}SS', & (89) \\ C &= M_{12}, & (90) \\ D &= M_{22} - M_{12}S. & (91) \end{aligned}$$

Now, the input and output rays are related to the system matrix M by Eqs. (57) and (58). Our aim is to use these equations to give a geometric interpretation to the M_{box} parameters M_{11}, M_{12}, M_{21}, and M_{22}. To do this, we shall choose from the set $\{x, x', \alpha, \alpha'\}$ a pair of variables and set the other variables constant. We shall consider four cases:

$$\begin{aligned} x' &= x'(x); & \alpha, \alpha' &= \text{constants} & (92) \\ x' &= x'(\alpha); & x, \alpha' &= \text{constants} & (93) \\ \alpha' &= \alpha'(x); & \alpha, x' &= \text{constants} & (94) \\ \alpha' &= \alpha'(\alpha); & x, x' &= \text{constants.} & (95) \end{aligned}$$

4.1 Telescopic System

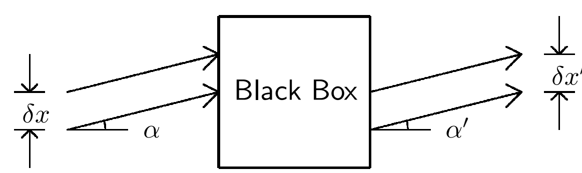

Figure 3: In a telescopic system, parallel rays emerge as parallel rays.

If $x' = x'(x)$, with the the angles α and α' constants, then the input and output beams are both bundles of parallel rays—a telescopic or afocal system. If this is true, then we may differentiate Eqs. (57) and (58) with respect to x to obtain[30, 31]

$$\begin{aligned} \frac{\partial x'}{\partial x} &= A &= M_{11} - M_{12}S', & \quad (96) \\ 0 &= C &= M_{12}. & \quad (97) \end{aligned}$$

Using these equations, together with Eqs. (57) to (58) and (88) to (91), we arrive at

$$\begin{aligned} \frac{\partial x'}{\partial x} &= M_{11}, & (98) \\ \frac{n'\alpha'}{n\alpha} &= M_{22}. & (99) \end{aligned}$$

Thus, in a telescopic system, the ratio of the outgoing and ingoing rays's index-dilated angles is the angular magnification M_{22}[32]; the ratio of the change in the output and input heights is M_{11}.

If we multiply Eqs. (98) and (99), we get

$$\frac{n'\alpha'}{n\alpha}\frac{\partial x'}{\partial x} = 1 = |\mathsf{M}_{box}|, \tag{100}$$

because $M_{12} = 0$. Equation (100) is the Lagrange's theorem for a telescopic system. An alternative formulation of this theorem is

$$n'\alpha'\,\delta x' = n\alpha\,\delta x, \tag{101}$$

where $n\alpha\,\delta x$ is the corresponding Lagrange invariant. That is, if the angles of the input and output rays are positive constants, then the change in height x of the input ray is proportional to the change in height of the output ray x'.

4.2 Inverse Fourier Transforming System

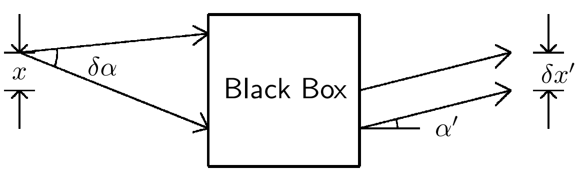

Figure 4: In an inverse Fourier transforming system, rays moving out from a point source become parallel rays.

If $x' = x'(\alpha)$, with x and α constants, then the system transforms a point source into a beam of paralell rays—an inverse Fourier transforming system. If this is true, then we may differentiate Eqs. (57) and (58) with respect to the index-dilated angle $n\alpha$ to obtain[30, 31]

$$\begin{aligned} \frac{\partial x'}{\partial(n\alpha)} &= B = \quad M_{21} + M_{22}S' \\ &\qquad\quad + M_{11}S - M_{12}SS', \qquad (102) \\ 0 &= D = \quad M_{22} - M_{12}S. \qquad (103) \end{aligned}$$

Using these equations, together with Eqs. (57) to (58) and (88) to (91), we arrive at

$$\begin{aligned} S &= \frac{M_{22}}{M_{12}} \equiv f, & (104) \\ \frac{n'\alpha'}{x} &= -M_{12}, & (105) \\ \frac{\partial x'}{\partial(n\alpha)} &= \frac{|\mathsf{M}_{\rm box}|}{M_{12}} = \frac{1}{M_{12}}. & (106) \end{aligned}$$

Thus, in an inverse Fourier transforming system, a point light source at the input focal distance $f = M_{22}/M_{12}$ to the left of the black box transforms to a bundle of parallel rays at the output[33]; the ratio of the optical inclination angle $n'\alpha'$ of the output beam to the height x of the point source is $-M_{12}$; the change in the height x' of the output ray with respect to the change in the dilated angle $n\alpha$ of the input ray is $1/M_{12}$.

If we multiply Eqs. (105) and (106), we get

$$\frac{n'\alpha'}{x}\frac{\partial x'}{\partial(n\alpha)} = -1. \tag{107}$$

Equation (107) is the Lagrange's theorem for an inverse Fourier transforming system. In terms of differentials, Eq. (107) may be written as

$$n'\alpha'\,\delta x' = -x\,\delta(n\alpha). \tag{108}$$

Notice that though we could not define a corresponding Lagrange invariant for this system, we can still provide a geometrical interpretation to Eq. (108): if the input height x and the output angle α' are positive constants, then a positive change in the angle α of the input ray would result to a negative change in the output height x' of the output ray.

4.3 Fourier Transforming System

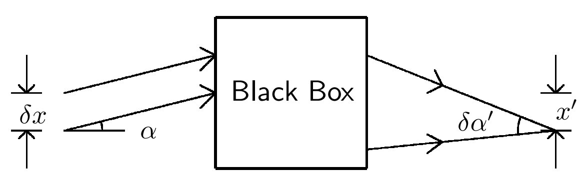

Figure 5: In a Fourier transforming system, parallel rays converge to a point. Note that the actual angle $\delta\alpha'$ is the angle opposite to the one labeled in the figure.

If $\alpha' = \alpha'(x)$, with x' and α constants, then the system focuses a beam of parallel rays into a point—a Fourier

transforming system. This means that we may differentiate Eqs. (57) and (58) with respect to x to obtain[30, 31]

$$
\begin{aligned}
0 &= A &= M_{11} - M_{12}S', \qquad (109)\\
\frac{\partial(n'\alpha')}{\partial x} &= -C &= -M_{12}. \qquad (110)
\end{aligned}
$$

Using these equations, together with Eqs. (57) to (58) and (88) to (91), we arrive at

$$
\begin{aligned}
S' &= \frac{M_{11}}{M_{12}} \equiv f', \qquad (111)\\
\frac{x'}{n\alpha} &= \frac{|\mathsf{M}_{\text{box}}|}{M_{12}} = \frac{1}{M_{12}}. \qquad (112)
\end{aligned}
$$

Notice that Eqs. (111), (112), and (110) are the conjugate relations for Eqs. (104) to (106). That is, in a Fourier transforming system, a bundle of parallel rays at the input is focused to a point source at the output at the focal distance $f' = M_{11}/M_{12}$ to the right of the black box; the ratio of the height x of the focus to the the dilated angle $n\alpha$ of the input beam is $-M_{12}$; the change in the dilated angle $n'\alpha'$ at the output focus with respect to the change in the height x of the input beam is $1/M_{12}$.

If we multiply Eqs. (110) and (112), we get

$$
\frac{x'}{n\alpha}\frac{\partial(n'\alpha')}{\partial x} = -1. \qquad (113)
$$

Equation (113) is the Lagrange's theorem for a Fourier transforming system. In terms of differentials, Eq. (113) may be written as

$$
x'\,\delta(n'\alpha') = -n\alpha\,\delta x. \qquad (114)
$$

Notice that though we could not also define a corresponding Lagrange invariant for this system, we could also still interpret Eq. (114) geometrically: if the input ray angle α and the output height x' are positive constants, then a positive change in the input ray height x would result to a negative change in the output ray angle α'.

4.4 Imaging System

If $\alpha' = \alpha'(\alpha)$, with x and x' constants, then the system transforms light from a point source to another point source—an imaging system. If this is true, then we may differentiate Eqs. (57) and (58) with respect to the input optical angle $n\alpha$ to obtain[30, 31]

$$
\begin{aligned}
0 &= B = M_{21} + M_{22}S'\\
&\qquad + M_{11}S - M_{12}SS', \qquad (115)\\
\frac{\partial(n'\alpha')}{\partial(n\alpha)} &= D = M_{22} - M_{12}S. \qquad (116)
\end{aligned}
$$

Using these equations, together with Eqs. (57) to (58) and (88) to (91), we arrive at[34, 35]

$$
\begin{aligned}
S' &= \frac{M_{11}S + M_{21}}{M_{12}S - M_{22}}, \qquad (117)\\
S &= \frac{M_{22}S' + M_{21}}{M_{12}S' - M_{11}}, \qquad (118)
\end{aligned}
$$

and

$$
\xi = \frac{x'}{x} = M_{11} - M_{12}S' = \frac{-1}{M_{12}S - M_{22}}, \qquad (119)
$$

where we used the identity $|\mathsf{M}_{\text{box}}| = 1$. Equation (117) are the Moebius relations for the object and image distances. Equation (119) is the definition for lateral magnification[36].

There are two relations that we can derive from these equations:

First, the product of Eqs. (116) and (119) is

$$
\frac{x'}{x}\frac{\partial(n'\alpha')}{\partial(n\alpha)} = -1, \qquad (120)
$$

which is the Lagrange's theorem for the imaging system. In terms of differentials, Eq. (120) may be written as

$$
x'\,\delta(n'\alpha') = -x\,\delta(n\alpha) \qquad (121)
$$

Notice that because of the presence of the negative sign, the quantity $x\delta(n\alpha)$ is not a Lagrange invariant; rather, it is its magnitude $|x\delta(n\alpha)|$. The geometrical interpretation of Eq. (121) is as follows: if the input and output ray heights are positive constants, then a positive change in the input angle α would result to a negative change in the output angle α'. (Note: the paraxial angle α is measured from the optical axis $\mathbf{e}_3$ pointing to the right. A positive α is a counterclockwise rotation; a negative α is clockwise. Thus, the actual $\delta\alpha'$ in Fig. 6 is the angle opposite to the one labeled in the figure, i.e., to the right of the image focus.)

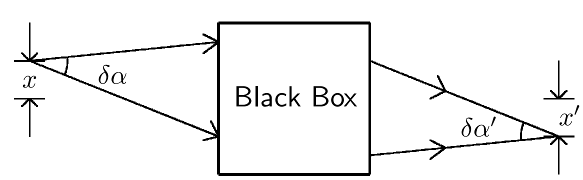

Figure 6: In an imaging system, rays leaving a point source converge to a point. Note that the actual angle $\delta\alpha'$ is the angle opposite to the one labeled in the figure.

Second, we may rewrite Eq. (119) as

$$(M_{12}S - M_{22})(M_{12}S' - M_{11}) = |\mathsf{M}_{\rm box}| = 1. \qquad (122)$$

In terms of the focal lengths f in Eq. (104) and f' in Eq. (111), Eq. (122) becomes

$$(S - f)(S' - f') = ZZ' = \frac{1}{M_{12}^2}. \qquad (123)$$

This is similar to the Newton's lens equation[38]

$$ZZ' = FF'. \qquad (124)$$

Note that the focal lengths f and f' are measured from the left-most and right-most side of the optical black box, while F and F' are measured from the left and right gaussian planes.

4.5 Summary

The system matrix M is defined as the ABCD matrix operating on input ray vector $\hat{r}$ to yield the output ray vector $\hat{r}'$

$$\begin{pmatrix} x' \\ in'\alpha' \end{pmatrix} = \begin{pmatrix} x \\ in\alpha \end{pmatrix} \begin{pmatrix} A & -iC \\ -iB & D \end{pmatrix}, \qquad (125)$$

where

$$A = \frac{\partial x'}{\partial x}, \qquad C = -\frac{\partial (n'\alpha')}{\partial x}, \qquad (126)$$

$$B = \frac{\partial x'}{\partial (n\alpha)}, \qquad D = \frac{\partial (n'\alpha')}{\partial (n\alpha)}. \qquad (127)$$

Because the system matrix has a unit determinant, then

$$1 = \frac{\partial x'}{\partial x}\frac{\partial (n'\alpha')}{\partial (n\alpha)} - \frac{\partial (n'\alpha')}{\partial x}\frac{\partial x'}{\partial (n\alpha)}, \qquad (128)$$

as given by Goodman.[37]

The four Lagrange theorems for the four optical system types may be combined into one:

$$\begin{aligned} 1 &= \frac{n'\alpha'}{n\alpha}\frac{\partial x'}{\partial x}; \qquad \alpha, \alpha' = const. \\ &= -\frac{n'\alpha'}{x}\frac{\partial x'}{\partial (n\alpha)}; \quad x, \alpha' = const. \\ &= -\frac{x'}{n\alpha}\frac{\partial (n'\alpha')}{\partial x}; \quad \alpha, x' = const. \\ &= -\frac{x'}{x}\frac{\partial (n'\alpha')}{\partial (n\alpha)}; \quad x, x' = const. \qquad (129) \end{aligned}$$

We shall refer to Eq. (129) as the unified Lagrange theorem for optical systems.

5 Conclusions

We used the orthonormality axiom in geometric algebra to show that the height-angle vector of a light ray must be complex. We proposed that the matrix operators to this vector should be right-acting, in order to follow the sequence of surfaces traversed by a light ray as it moves from left-to-right close to the optical axis. We showed that the propagation and refraction matrix operators may be expressed as a sum of a unit matrix and a imaginary non-diagonal Fermion matrix. We developed combinatorial rules for finding the product of a succession of propagation and refraction matrices, without doing explicit matrix multiplication. This product is the right-acting system matrix with real and imaginary coefficients.

We factored out the system matrix as a product of the input propagation matrix, the black box matrix, and the output propagation matrix. Based on the coefficients of the system matrix, we classified the optical systems into four: telescopic, inverse Fourier transforming, Fourier transforming, and imaging. We showed that all four systems have a corresponding Lagrange theorem expressed in partial derivatives, and that these theorems may be combined into one. We transformed these theorems in terms of Lagrange differentials, which allows us to geometrically interpret the effect of the change in input variable to the change in the output variable. We showed that these differential relations result to the Lagrange invariants only for the telescopic and imaging systems.

In a future work, we shall revisit paraxial skew ray tracing using complex vectors and right-acting matrices.

Acknowledgments

This research was supported by the Manila Observatory and by the Physics Department of Ateneo de Manila University.

References

[1] Allen Nussbaum and Richard A. Phillips, *Contemporary Optics for Scientists and Engineers* (Prentice-Hall, Englewood Cliffs, NJ, 1976), p. 10.

[2] William E. Baylis, J. Huschilt, and Jiansu Wei, "Why i?" Am. J. Phys. **60**(9), 788–797 (1992). See p. 789. Baylis et al's form is $\mathbf{e}_j\mathbf{e}_k + \mathbf{e}_k\mathbf{e}_j = 2\delta_{jk}$.

[3] Amnon Yariv, *Quantum Electronics* (John Wiley, New York, 1989), pp. 106–107.

[4] Bernard Jancewicz, *Multivectors and Clifford Algebra in Electrodynamics* (World Scientific, Singapore, 1988), p. 32.

[5] David Hestenes, "Oersted medal lecture 2002: Reforming the mathematical language of physics," Am. J. Phys. **71**(2), 104–121 (2003). See p. 110.

[6] Quirino M. Sugon Jr., Carlo B. Fernandez, and Daniel J. McNamara, "A geometric algebra reformulation of 2×2 matrices: the dihedral group $\mathcal{D}_4$ in bra-ket notation," arXiv:0811.3680v1 [math-ph] (24 Nov 2008). See p. 9.

[7] Keith R. Symon, *Mechanics* (Addison-Wesley, Reading, MA, 1971), 3^{rd} ed., p. 408.

[8] Charles E. Campbell, "Ray vector fields," J. Opt. Soc. Am. A **11**(2), 618–622 (1994).

[9] W. F. Harris, "Dioptric power: its nature and its representation in three- and four-dimensional space," Opt. Vis. Sci., **74**(6), 349–366 (1997). See p. 357.

[10] Jun John Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, MA, 1967), p. 28.

[11] Michel Le Bellac, *Quantum and Statistical Field Theory* (Clarendon, Oxford, 1991), p. 404.

[12] Miles V. Klein and Thomas E. Furtak, *Optics*, (John Wiley, New York, 1986), 2^{nd} ed., p. 152.

[13] Jurgen R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 2nd ed. (Prentice-Hall, Englewood Cliffs, NJ, 1984), p. 55.

[14] Max Born and Emil Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Pergamon, Oxford, 1964), p. 165.

[15] W. T. Welford, *Aberrations of the Optical System* (Academic, London, 1974), pp. 22–23.

[16] See Ref. [6], p. 8.

[17] Quirino M. Sugon Jr. and Daniel J. McNamara, "A geometric algebra reformulation of geometric optics," Am. J. Phys. **72**(1), 92-97. See p. 95. The concavity sign function $c_{\sigma\eta}$ is defined as the normalized dot product of the incident ray $\boldsymbol{\sigma}$ and the normal vector $\boldsymbol{\eta}$.

[18] Quirino M. Sugon Jr. and Daniel J. McNamara, "Ray tracing in spherical interfaces using geometric algebra," in *Advances in Imaging and Electron Physics* **139**, 179–224. See pp. 206. The concavity function $c_{\sigma\eta}$ is generalized to the relative direction sign function c_{vw} for two vectors $\mathbf{v}$ and $\mathbf{w}$.

[19] Quirino M. Sugon Jr. and Daniel J. McNamara, "Paraxial meridional ray tracing equations from the unified reflection-refraction law via geometric algebra," arXiv:0810.5224v1 [physics.optics] (29 Oct 2008). See p. 3.

[20] Eugene Hecht, *Optics*, 2nd ed., with contributions by Alfred Zajac (Addison Wesley, Reading, MA, 1987), p. 217.

[21] See Ref. [19], p. 4.

[22] R. Simon and Kurt Bernardo Wolf, "Structure of the set of paraxial optical systems," J. Opt. Soc. Am. **17**(2), 342-354 (2000). See pp. 345-347.

[23] See Ref. [20], p. 219.

[24] J. Warren Blaker and William M. Rosenblum, *Optics: An Introduction for Students of Engineering* (MacMillan, New York, 1993), p. 37.

[25] See Ref. [12], p. 156.

[26] See Ref. [12], p. 157.

[27] Joseph W. Simmons and Mark J. Guttmann, *States, Waves and Photons: A Modern Introduction to Light* (Addison-Wesley, Reading, MA, 1970), p. 8.

[28] Ajoy Ghatak, *Optics* (Tata, New Delhi, 1977), p. 52.

[29] See Ref. [1], p. 52.

[30] Douglas S. Goodman, "General Principles of Geometric Optics," in *Handbook of Optics*, chapter 1 of vol. 1, *Fundamentals, Techniques, and Design* (McGraw-Hill, New York, 1995), p. 71.

[31] Moshe Nazarathy and Joseph Shamir, "First-order optics—a canonical operator representation: lossless systems," J. Opt. Soc. Am. **72**(3), 356–364 (1982). See p. 362.

[32] See Ref. [27], p. 11.

[33] See Ref. [27], p. 10.

[34] See Ref. [1], p. 16.

[35] See Ref. [28], p. 56.

[36] See Ref. [27], p. 12.

[37] See Ref. [30], p. 45.

[38] See Ref. [13], p. 48.